REMARKS

I. Introduction

In response to the Office Action dated December 17, 2002, no claims have been cancelled, amended or added. Claims 1-24 remain in the application. Re-examination and re-consideration of the application, as amended, is requested.

II. Specification Amendments

Applicants' attorney has made amendments to the specification as indicated above, in response to the request set forth in the Office Action.

III. Prior Art Rejections

A. The Office Action Rejections

In paragraphs (3)-(4) of the Office Action, claims 1, 3, 7, 9, 11, 15, 17, 19, and 23 were rejected under 35 U.S.C. §102(e) as being anticipated by Fayyad et al., U.S. Patent No. 6,263,337 (Fayyad). In paragraphs (5)-(6) of the Office Action, claims 2, 4-6, 12-14, 18, and 20-22 were rejected under 35 U.S.C. §103(a) as being unpatentable over Fayyad in view of Van Huben et al., U.S. Patent No. 6,327,594 (Van Huben). In paragraph (7) of the Office Action, claims 8, 16, and 24 were rejected under Fayyad in view of Guha et al., U.S. Patent No. 6,049,797 (Guha).

Applicants' attorney respectfully traverses these rejections.

B. The Applicants' Independent Claims

Independent claim 1 is directed to a data structure for analyzing data in a computer-implemented data mining system, wherein the data structure is a data model that comprises a Gaussian Mixture Model that stores transactional data, and the data model is mapped to aggregate the transactional data for cluster analysis.

Independent claim 9 is directed to a method for analyzing data in a computer-implemented data mining system, comprising:

generating a data structure in the computer-implemented data mining system, wherein the data structure is a data model that comprises a Gaussian Mixture Model that stores transactional data; and

mapping the data model to aggregate the transactional data for cluster analysis.

Independent claim 17 is directed to an apparatus for analyzing data in a computer-implemented data mining system, comprising:

means for generating a data structure in the computer-implemented data mining system, wherein the data structure is a data model that comprises a Gaussian Mixture Model that stores transactional data; and

means for mapping the data model to aggregate the transactional data for cluster analysis..

C. The Fayyad Reference

Fayyad describes one exemplary embodiment providing a data mining system for use in finding clusters of data items in a database or any other data storage medium. Before the data evaluation begins a choice is made of the number M of models to be explored, and the number of clusters (K) of clusters within each of the M models. The clusters are used in categorizing the data in the database into K different clusters within each model. An initial set of estimates for a data distribution of each model to be explored is provided. Then a portion of the data in the database is read from a storage medium and brought into a rapid access memory buffer whose size is determined by the user or operating system depending on available memory resources. Data contained in the data buffer is used to update the original model data disttibutions in each of the K clusters over all M models. Some of the data belonging to a cluster is summarized or compressed and stored as a reduced form of the data representing sufficient statistics of the data. More data is accessed from the database and the models are updated. An updated set of parameters for the clusters is determined from the summarized data (sufficient statistics) and the newly acquired data. Stopping criteria are evaluated to determine if further data should be accessed from the database.

D. The Van Huben Reference

Van Huben describes a common access method to enable disparate pervasive computing devices to interact with centralized data management systems. A modular, scalable data management system is envisioned to further expand the role of the pervasive devices as direct participants in the data management system. This data management system has a plurality of data managers and is provided with a plurality of data managers in one or more layers of a layered architecture. The system performs with a data manager and with a input from a user or pervasive computing device via an API a plurality of process on data residing in heterogeneous data repositories of computer system including promotion, check-in, check-out, locking, library searching, setting and viewing

process results, tracking aggregations, and managing parts, releases and problem fix data under management control of a virtual control repository having one or more physical heterogeneous repositories. The system provides for storing, accessing, tracking data residing in said one or more data repositories managed by the virtual control repository. DMS applications executing directly within, on or behalf of, the pervasive computing device organize data using the PFVL paradigm. Configurable managers include a query control repository for existence of peer managers and provide logic switches to dynamically interact with peers. A control repository layer provides a common process interface across all managers. A command translator performs the appropriate mapping of generic control repository layer calls to the required function for the underlying storage engine.

E. The Guha Reference

Guha describes an invention relating to a computer method, apparatus and programmed medium for clustering databases containing data with categorical attributes. The present invention assigns a pair of points to be neighbors if their similarity exceeds a certain threshold. The similarity value for pairs of points can be based on non-metric information. The present invention determines a total number of links between each cluster and every other cluster bases upon the neighbors of the clusters. A goodness measure between each cluster and every other cluster based upon the total number of links between each cluster and every other cluster and the total number of points within each cluster and every other cluster is then calculated. The present invention merges the two clusters with the best goodness measure. Thus, clustering is performed accurately and efficiently by merging data based on the amount of links between the data to be clustered.

F. The Applicants' Claims Are Patentable Over The References

Applicants' invention, as recited in independent claims 1, 9 and 17, is patentable over the references, because the claims recite limitations not found in the references. Specifically, Fayyad does not disclose that the data model is mapped to aggregate the transactional data for cluster analysis.

The Office Action cites Fayyad as teaching all the elements of the independent claims, including a data structure for analyzing data in a computer-implemented data mining system, as reference number 12 in FIG. 2 and in the accompanying text. The Office Action also cites Fayyad as teaching that the data structure is a data model that comprises a Gaussian Mixture Model that

stores transactional data, at col. 9, lines 22-67. In addition, the Office Action cites Fayyad as teaching that the data model is mapped to aggregate the transactional data for cluster analysis, at col. 8, lines 34-46. With regard to the method and apparatus claims, the Office Action cites Fayyad as teaching generating the data structure referred to above, at col. 9, line 57 to col. 11, line 29.

Applicants' attorney disagrees. At the location indicated above, Fayyad does not teach that the data model is mapped to aggregate the transactional data for cluster analysis:

> Fayyad: col. 8, lines 34-46
> An additional data structure designated DS in FIG. 6A includes an array of pointers 160 that point to a group of k vectors (the cluster number) of n elements 162 designated 'SUM' a second group of k vectors 164 designated 'SUMSQ', and a group 166 of k floats designated M. This data structure is similar to the data structure of FIG. 6D that describes the MODEL. It contains sufficient statistics for a number of data records that have been compressed into the data structure shown rather than maintained in memory. Compression of the data into this data structure and the CS data structure described below frees up memory for accessing other data from the database at the step 10 on a next subsequent iteration of the FIG. 4 scalable EM process.

The above description in Fayyad merely describes an array of pointers to summation values, which are statistics for records that have been compressed. However, it does not describe mapping the data model to aggregate transactional data for cluster analysis.

Moreover, Van Huben and Guha fail to overcome these limitations of Fayyad. Recall that Van Huben and Guha were only cited only against some of the dependent claims.

Thus, the references do not teach or suggest Applicants' invention. Moreover, the various elements of Applicants' claimed invention together provide operational advantages over the references. In addition, Applicants' invention solves problems not recognized by the references.

Thus, Applicants' attorney submits that independent claims 1, 9 and 17 are allowable over the references. Further, dependent claims 2-8, 10-16 and 18-24 are submitted to be allowable over the references in the same manner, because they are dependent on independent claims 1, 9 and 17, respectively, and thus contain all the limitations of the independent claims. In addition, dependent claims 2-8, 10-16 and 18-24 recite additional novel elements not shown by the references.

IV. Conclusion

In view of the above, it is submitted that this application is now in good order for allowance and such allowance is respectfully solicited. Should the Examiner believe minor matters still remain that

can be resolved in a telephone interview, the Examiner is urged to call Applicants' undersigned attorney.

Respectfully submitted,

GATES & COOPER LLP
Attorneys for Applicants

Howard Hughes Center
6701 Center Drive West, Suite 1050
Los Angeles, California 90045
(310) 641-8797

By: _____
Name: George H. Gates
Reg. No.: 33,500

Date: March 17, 2003

GHG/sjm/amb

G&C 30145.408-US-01